Dynamic Channel Management in MIMO OFDM Cellular Systems

Ben Lu ∗ Xiaodong Wang † Richard D. Gitlin ∗
Mohammad Madihian ∗

Abstract

We consider dynamic channel management (DCM), more specifically power control and channel allocation, in multi-carrier multi-antenna wireless cellular systems. A quasi-static channel model is assumed such that the channels remain approximately time-invariant within each packet. On a packet-by-packet basis, the mobile station (MS) receivers feed back the DCM parameters to the corresponding base station (BS) transmitters in order to minimize the transmission power and strive to satisfy every individual user’s data-rate request. While optimum, centralized DCM for all active cells typically requires solution of a non-convex optimization problem (which is subject to local minima) and requires intense channel feedback to a central controller. We instead consider sub-optimal distributed DCM, where each individual cell locally optimizes its own performance based on the local channel information. In particular, we compare three distributed DCM schemes, namely, the two-dimensional DCM which jointly adjusts both spatial and frequency sub-channels of each BS-MS link, the one-dimensional DCM which adjusts only the frequency sub-channels and the scalar DCM which adjusts each BS-MS link by one scalar power coefficient. These DCM schemes are investigated in multi-cell environments with realistic parameters; and it is seen that due to the multiple antennas, the maximum supportable rate (i.e., the rate supported by practically finite transmission power) is significantly increased, e.g., by a factor of two for users using four antennas and close to cell boundary.

Index Terms: dynamic channel management, multi-carrier OFDM, multi-antenna MIMO, wireless cellular systems, convex optimization

∗NEC Laboratories America, Princeton, NJ 08540.
†Columbia University, New York, NY 10027.
1 Introduction

Efficient power control and channel allocation are crucial to reduce the interference and to achieve the quality of service (QoS) goals in wireless cellular networks. In traditional narrow-band wireless networks, various power control algorithms have been studied in the last decade, e.g., [1, 2, 3, 4, 5]. Recently, multi-carrier and multi-antenna (i.e., MIMO) techniques have emerged as promising candidate solutions for high-data-rate wireless services. Multi-carrier modulation is often referred as discrete multi-tone (DMT) modulation in digital subscriber loop (DSL) systems; where the power control and bit-loading algorithms have been extensively investigated in [6, 7, 8, 9]. In wireless communications, multi-carrier modulation typically refers to orthogonal frequency-division multiplexing (OFDM); where many schemes of resource allocations (e.g., power control, sub-carrier channel allocation and bit-loading) have been explored for multiuser OFDM systems, e.g., [10, 11, 12, 13, 14]. On the other hand, research on the different aspects of MIMO techniques has been notably active in the past a few years; in particular, for narrow-band cellular MIMO systems, the attainable data throughput under maximal transmission power constrain was studied in [15], and the impact of multi-cell interference on MIMO capacity was explored in [16, 17].

When put together, multi-carrier (OFDM) and multi-antenna (MIMO) techniques result in a bundle of sub-channels both in frequency and in spatial domain. In this work, we restrict our attention to MIMO-OFDM downlink transmission in multi-cell systems; although essentially the same framework applies to uplink transmission as well. A low mobility model is assumed such that all downlink channels are slowly time-varying and remain approximately constant during the interval of each packet (which consists of a number of OFDM words). On a packet-by-packet basis, the receivers feed back the dynamic channel management (DCM) parameters to the corresponding transmitters in order to minimize the transmission power and to meet each individual user’s data-rate request.

In approaching the problem, we first formulate the DCM problem as minimizing the total transmission power of all base stations while satisfying each user’s data rate request. However, such a global DCM problem in general is a non-convex optimization problem, meaning its solution is subject to local minima; and it usually requires intense channel feedback to a central controller. We instead turn to a sub-optimal approach, in which each cell locally carries out the DCM based solely on the local channel information and by treating inter-cell interference as noise. In particular, we derive optimal solutions to three different DCM schemes from the basic convex optimization
theory; and it turns out that the solutions bear the form of the well-known “water-filling” strategy. More specifically, these DCM schemes are the two-dimensional DCM which adjusts both spatial and frequency sub-channels of each BS-MS link, the one-dimensional DCM which adjusts only the frequency sub-channels and the scalar DCM which adjusts each BS-MS link by one scalar power coefficient. The two-dimensional DCM provides the finest channel management, and its performance gain over the other two DCM schemes is larger especially at lower data rates and/or with the use of more antennas. On the other hand, the one-dimensional DCM is more capable of taking advantage of the asymmetry in sub-carrier channels than the scalar DCM; but such a performance advantage is considerably reduced when more antennas are employed by the system, because the variation of achievable data-rate on different sub-carriers is averaged down by the multiple spatial-channels. All the above DCM schemes are simulated in multi-cell environments with realistic parameters; and significant performance gain due to the employment of multiple antennas is exhibited.

The rest of this paper is organized as follows. In Section 2, the multi-cell MIMO-OFDM system model is given. In Section 3, both centralized and distributed DCA problems are formulated; and optimal solutions to distributed DCA problems are derived. In Section 4, the performance of the distributed DCA schemes is shown by numerical simulations in both single-cell and multi-cell systems. Finally, Section 5 contains the conclusions.

2 System Model

We consider the downlink transmission in a multi-carrier multi-antenna cellular system and assume that the intra-cell users share the channel by time-division multiple-access (TDMA) and the frequency reuse factor across multiple cells is one. In practice, due to the propagation delay, the intra-cell received signals and the inter-cell interfering signals are usually asynchronous; for simplicity, we neglect this asynchrony and adopt a synchronous signal model. \(^1\)

System-wide, at each moment there are \(J\) BS’s communicating to \(J\) MS’s; a scenario known as interference channels in the literature [18]. For a particular OFDM word within each packet, the received signal at the \(k\)-th frequency sub-carrier of the \(j\)-th mobile station (MS) (i.e., the MS in

\(^1\)On the other hand, the asynchronous inter-cell packets can be better approximated (by the central limit theorem) as non-cooperative Gaussian processes, which is one of underlying assumptions in Eqn. (2).
the $j$-th cell), $j = 1, \ldots, J$, can be written as

$$y_{j,k} = \sqrt{\beta_{j,j} d_{j,j}^{-\gamma}} \cdot H_{j,j,k} Q_{j,j,k} \sqrt{P_{j,k}} x_{j,k} + \sum_{i \neq j} \sqrt{\beta_{i,j} d_{i,j}^{-\gamma}} \cdot H_{i,j,k} Q_{i,j,k} \sqrt{P_{i,k}} x_{i,k} + n_{j,k},$$

$k = 1, \ldots, K,$

where

- $K$ is the number of total frequency sub-carriers; $N$ is the number of transmit antennas; $M$ is the number of receive antennas; and $J$ is the number of cells in the system.

- $y_{j,k}$ is an $M$-dimensional received signal vector of the $j$-th MS at the $k$-th frequency subcarrier. Similarly, $x_{j,k}$ is the $N$-dimensional transmit signal vector; and $n_{j,k}$ is the $N$-dimensional circularly symmetrical white Gaussian noise vector with zero mean and identity covariance matrix.

- $\beta_{i,j}$ is a log-normal shadowing fading variable from the $i$-th base station (BS) to the $j$-th MS, i.e., $10 \log_{10} \beta_{i,j}$ is a zero-mean Gaussian random variable with standard deviation $\sigma$; it is assumed that $\beta_{i,j}$ is location specific and takes the same value for all subcarriers. In addition, $d_{i,j}$ is the propagation distance between the $i$-th BS and the $j$-th MS; and $\gamma$ is the path loss exponent.

- $H_{i,j,k}$ is an $(M \times N)$ Rayleigh fading channel matrix at the $k$-th subcarrier for the air-link from the $i$-th BS to the $j$-th MS. With the proper insertion of cyclic prefix in every OFDM word, $H_{i,j,k}$ is determined by the discrete fourier transform (DFT) of the time-domain multi-path response matrices $h_{i,j,l}$, $l = 1, \ldots, L$, where $h_{i,j,l}$ is an $M \times N$ matrix consisting of circularly symmetrical Gaussian random variables with unit variance and zero means. $H_{i,j,k}$ at different sub-carriers (i.e., for different $k$) are generally correlated.

- $Q_{i,j,k}$ is an $(M \times M)$ precoding matrix for the air-link from the $i$-th BS to the $j$-th MS. $P_{j,k}$ is an $(M \times M)$ diagonal matrix of transmission power allocation of the $j$-th cell.

Some system assumptions in our study are as follows. It is assumed that all the mobile users in the systems are moving slowly, so that the channels remain approximately quasi-static during the interval of one packet. On a packet-by-packet basis, the MS’s compute and feedback the DCM parameters to the corresponding BS’s through dedicated uplink control channels. For simplicity, we have omitted the dynamics of data traffic and further assumed that the $J$ cells are active all
the time. It is also assumed that all BS’s transmit properly designed pilot symbols in each packet, such that the MSs are able to estimate the downlink channels (including the desired one and the interfering ones) with sufficient precision;\(^2\) whereas the assumption of perfect channel information at the BS transmitter side is required or not by particular DCM scheme, as will be explained in next section.

For better understanding, an exemplary time basis frame structure of a particular BS-MS link in the considered cellular system is illustrated in Fig 1. Each downlink packet starts with the broadcast of pilot sequences that received by all MSs; and the feedback of DCM parameters (calculated at the MSs) are sent to the BS through the uplink channels; upon receiving of those DCM parameters, certain dynamic scheduling is performed at the BS; then in the remainder of the packet (for the duration of a number of OFDM words), the BS communicates to that scheduled MS exclusively.

A further investigation of the dynamic scheduling algorithm is out of the scope of this paper; among other things, we note that such a scheduling scheme is differentiated from the existing ones (e.g., [19]) in the sense that multiuser diversity [20] is utilized to minimize the transmission power (hence reduce interference) in support of the pre-determined data rates of MSs. The downlink and uplink coexist either by time-division duplexing (TDD) or frequency-division duplexing (FDD); in addition, in the description above, we implicitly assumed a cell-by-cell distributed control, the justification to which will be given in Section 3.

---

\(^2\)Note that estimating all the downlink channels, seemingly demanding though, actually appears in practical systems, e.g., in cdmaOne system (mostly for soft-handoff purpose) the MS periodically screens/estimates neighboring cells’ downlink channels.
3 Dynamic Channel Management

The basic problem of interest is how to satisfy individual user’s quality of service (QoS) with the minimum expense of system resources.

Specifically, we consider the system resources as the transmission power and the physical channels. Transmitting no power beyond that is required by the QoS is very crucial to reduce the interference in wireless networks (even though the power-saving may not be a priority concern for base stations). In a multi-carrier multi-antenna system, the physical channels can be decomposed into a bundle of sub-channels, each corresponding to a particular frequency sub-carrier, time slot and spatial sub-channel. As illustrated in Fig 1, the proper dynamic scheduling of time-slots among intra-cell users can reduce the transmission power to support a given data rate by taking advantage of the multiuser diversity, a concept introduced in [20]. Once the time-slots are assigned to different intra-cell users, each user is left with two-dimensional sub-channels in both frequency and spatial domains.

In single-antenna systems, the single-to-interference-plus-noise-ratio (SINR) is often adopted as the QoS criterion, e.g., in [11, 5]. As it is not straightforward to define the SINR in multi-antenna systems, we adopt the information rate as the QoS criterion. For each user, the information rate is the aggregate mutual information at all frequency sub-carriers with MIMO. By treating the inter-cell interference and ambient noise as non-cooperative noise (i.e., not trying to decode inter-cell signals), the mutual information (in the unit of nats/second) at the $k$-th sub-carrier of the $j$-th MS is given by [16]

$$I_{j,k} = \log \det \left[ I_M + \left( I_M + \sum_{i \neq j} \beta_{i,j} d_{i,j}^{-\gamma} H_{i,j,k} Q_{i,j,k} P_{i,k} Q_{i,j,k}^H H_{i,j,k}^H \right)^{-1} \right]$$

$$= \log \det \left[ \beta_{j,j} d_{j,j}^{-\gamma} H_{j,j,k} Q_{j,j,k} P_{j,k} Q_{j,j,k}^H H_{j,j,k}^H \right].$$

The information rate can be achieved by practical coding and modulation schemes with a small SNR gap. Given the channel state information, a matrix singular-value-decomposition (SVD) can be computed as $R_{j,k}^{-1/2} H_{j,j,k} = U_{j,k} \Phi_{j,k} V_{j,k}$, where $U_{j,k}$ and $V_{j,k}$ are unitary matrices and $\Phi_{j,k}$ is a diagonal matrix. When perfect channel information (of both the desired channel and interfering channels) is available at the base station, it is optimal to use the precoding matrix $Q_{j,j,k} = V_{j,k}^H$. 

[16]; then (2) can be simplified as

$$I_{j,k} = \log \det \left( I_M + \beta_{j,j} d_{j,j}^{-1} \Phi_{j,k} P_{j,k} \Phi_{j,k} \right) = \sum_{s=1}^{m} \frac{\log(1 + P_{j,k,s} \phi_{j,k,s})}{I_{j,k,s}} \tag{3}$$

where \( m \triangleq \min(N, M) \); \( \phi_{j,k,s}, \forall s \), are the diagonal elements of nonnegative diagonal matrix \( (\beta_{j,j} d_{j,j}^{-1} \Phi_{j,k} \Phi_{j,k}) \); and \( I_{j,k,s} \) denotes the mutual information of the \( k \)-th sub-carrier and the \( s \)-th spatial sub-channel of the \( j \)-th MS. It is assumed that \( \Phi_{j,k} \neq 0 \), \( \forall k \); otherwise that sub-carrier MIMO channel is excluded from consideration. Without channel information at the base station, it is optimal to let \( Q_{j,j,k} = I_N \) and \( P_{j,k} = P_{j,k} I_N \) [21]; then (2) can be simplified as

$$I_{j,k} = \sum_{s=1}^{m} \frac{\log(1 + P_{i,k,s} \phi_{i,k,s})}{I_{i,k,s}} \tag{4}$$

where \( \phi_{i,k,s} \) is the same as that in (3). The only difference in (4) is that power control coefficient \( P_{i,k} \) takes the same value for all spatial sub-channels at the \( k \)-th sub-carrier. Because without transmitter precoding, the spatial sub-channels are not differentiable from each other; and finer channel management on spatial-sub-channel level becomes impossible.

### 3.1 Centralized Dynamic Channel Management

First, we formulate the problem as minimizing the total transmission power of all \( J \) cells, and simultaneously satisfying all users’ data rate goals. That is,

$$\min \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{s=1}^{m} \mu_{j,k,s} P_{j,k,s} (\{I_{j,k,s}\}) \tag{5}$$

subject to

$$\sum_{k=1}^{K} \sum_{s=1}^{m} \mu_{j,k,s} I_{j,k,s} \geq R_j, \quad j = 1, 2, \ldots, J,$$

$$\mu_{j,k,s} (\mu_{j,k,s} - 1) = 0, \quad \forall j, \forall k, \forall s,$$

$$I_{j,k,s} \geq 0, \quad \forall k, \forall s$$

where \( R_j \) is the requested data rate of the \( j \)-th MS; \( \mu_{j,k,s} \) is the channel allocation function, taking value of 1 or 0; without loss of clarity, the same notation \( P_{j,k,s} (\{I_{j,k,s}\}) \) represents either the function or the function value of \( P_{j,k,s} \) with respect to \( \{I_{j,k,s}\} \). The objective is to optimize \( P_{j,k,s} \) and \( \mu_{j,k,s} \) in order to minimize the total transmission power while satisfy each user’s data rate request. However, the problem in (5) is in general a non-convex problem, which means its
solution is subject to local minima. Moreover, the intense channel information feedback to a central controller is required; note that it is costly as well to implement such a scheme in TDD systems, as the reciprocity between the desired downlink and uplink channels is not applicable to inter-cell interfering channels.

3.2 Distributed Dynamic Channel Management

A lower-complexity sub-optimal approach is to solve the global DCM problem in (5) by solving a series of distributed DCM problems. In particular, we can minimize the total transmission power in the $j$-th cell, subject to the requirement for this cell, i.e.,

$$\min \sum_{k=1}^{K} \sum_{s=1}^{N} P_{j,k,s}(\{I_{j,k,s}\}_{k,s}),$$

(6)

subject to

$$\sum_{k=1}^{K} \sum_{s=1}^{m} I_{j,k,s} \geq R_{j},$$

(7)

$$I_{j,k,s} \geq 0, \forall k, \forall s.$$  

(8)

It is worth to note that compared to (5), the channel allocation parameters $\mu_{j,k,s}$ are dropped in (6). It can be shown that (6) is a standard convex optimization problem, and the water-filling power allocation of $P_{j,k,s}$ gives the optimal solution. Therefore, without loss of optimality, a sub-channel is allocated if its optimal power allocation $P^{*}_{j,k,s} > 0$ and it is not allocated if $P^{*}_{j,k,s} = 0$. (The superscript * denotes the optimal value of individual variable.) This intuitive argument is made rigorous in the Appendix.

Once each individual cell obtains the optimal solution to the local problem (6), the solution to the global problem (5) is computed in an iterative way as

for $r=1$ to $t$

for $j=1$ to $J$

Find the optimal solution to the local DCM problem of the $j$-th cell, by solving (6)-(8).

end

end

where $t$ is a small integer; since the local DCM of individual cell will change the interference profile to other cells, a total of $t$ iterations is implemented for all $J$ cells to achieve equilibrium, although such a solution may only be a local minimum of (5). We observed in numerical experiments that the above iterative algorithm typically converges in less than $t = 3 \sim 4$ iterations. In practice,
because of the path attenuation, the iteration may take place among small groups of neighboring
cells. Moreover, rather than simultaneously updating power allocations of all cells (where the rate
constraint could be violated if the inter-cell interference profile is concurrently changing), different
cells can sequentially carry out the distributed DCM, e.g., as the staggered DCA proposed in [11].
We in particular give credit to a previous work studying the distributed power control of multiuser
DSL systems [9], in which a similar distributed power control scheme was interpreted from the
viewpoint of non-cooperative games.

In what follows, we present three distributed DCM algorithms working under different system
assumptions; for notational convenience, the subscript \( j \) that denotes the \( j \)-th cell will be dropped.

### 3.2.1 Two-dimensional Distributed DCM Algorithm

Let’s consider the distributed DCM algorithm, when perfect channel information is available at the transmitter. First, by the cyclic prefix insertion, OFDM modulation leads to orthogonal frequency-domain sub-channels. Moreover, when transmitter precoding is carried out as in (3), the orthogonality of spatial-domain sub-channels can also be achieved. Overall, for each physical link as defined in (1), there are as many as \( mK \) sub-channels; and the goal of the two-dimensional distributed DCM is to optimally control the power and allocate the sub-channels in both frequency and spatial domains. Given the particular channel realization \( H_{k,s}, \forall k, \forall s \), these \( mK \) sub-channels are mutually orthogonal; therefore, the two-dimensional distributed DCM problem degenerates into the traditional parallel Gaussian channel capacity problem, where the total \( mK \) sub-channels can be considered independently. It is known from [22, 23] that the water-filling algorithm is the optimal solution to such problems. The two-dimensional distributed DCM is equivalent to numerically solving the following equations,

\[
P_{k,s}^* = \left[ \frac{\lambda^* - 1}{\phi_{k,s}} \right]^+ , \quad \forall k, \forall s ,
\]

\[
\sum_{k=1}^{K} \sum_{s=1}^{m} I_{k,s}(P_{k,s}^*) = R ,
\]

where \( [x]^+ = x \), if \( x > 0 \); \( [x]^+ = 0 \), if \( x \leq 0 \); and \( I_{k,s}(P_{k,s}) = \log(1 + P_{k,s}\phi_{k,s}) \). While the two-dimensional DCM algorithm provides the finest control over sub-channels, its requirement of ideal channel information at the transmitter side (for transmitter precoding) and the computation of matrix SVD constitute a major obstacle in practical applications. We next consider the schemes without transmitter precoding.
3.2.2 One-dimensional Distributed DCM Algorithm

The one-dimensional distributed DCM is formulated as

\[
\min \sum_{k=1}^{K} N \cdot P_k(I_k),
\]

subject to

\[
\sum_{k=1}^{K} I_k \geq R,
\]

\[
I_k \geq 0, \forall k,
\]

which is slightly different from the generic formulation in (6). Because, \textit{when the channel information is not available at the transmitter}, by (4) it is optimal to transmit the same power \(P_k\) from all \(N\) transmit antennas, and it is more convenient to consider the collective information rate \(I_k\), on the \(k\)-th sub-carrier. It is shown in the Appendix that the problem in (11) is a standard convex optimization problem, and its solution is the generalized water-filling by solving the following equations,

\[
P_k^* = P_k(g^{(-1)}(I_k)) = \lambda^*, \forall k, \quad \text{if} \quad \lambda^* > \left[ \sum_{s=1}^{m} \phi_{k,s} \right]^{-1},
\]

\[
P_k^* = 0, \quad \forall k, \quad \text{if} \quad \lambda^* \leq \left[ \sum_{s=1}^{m} \phi_{k,s} \right]^{-1},
\]

\[
\sum_{k=1}^{K} I_k(P_k^*) = R,
\]

where \(g(I_k) \triangleq \frac{dP_k(I_k)}{dI_k}\), the derivative of \(P_k\) with respect to \(I_k\); \(g^{(-1)}(I_k)\) denotes the inverse function of \(g(I_k)\); and \(I_k(P_k)\) is a single-variable function of \(P_k\) as defined in (4).

In general, (14)-(16) can be solved using simple numerical techniques. For a few simple cases, we are able to obtain the analytical expression for the function \(g^{(-1)}(I_k)\), and further reduce the implementation complexity in solving (14)-(16).

**Special Case: Single-antenna Systems \((N = 1, M = 1)\)**

From (4), we have

\[
g^{(-1)}(I_k) = \log(g(I_k) \cdot \phi_{k,1}).
\]

Hence, the optimal \(P_k^*\) is given as

\[
P_k^* = [\lambda^* - 1/\phi_{k,1}]^+,
\]

where \(\lambda^*\) is further determined by the rate condition (16).
Special Case: Dual-Antenna Systems \((N = 2, M = 2)\)

After some simple algebra on (4), we get
\[
g(I_k) = \frac{e^{I_k}}{\sqrt{(\phi_{k,1} + \phi_{k,2})^2 + 4\phi_{k,1}\phi_{k,2}(e^{I_k} - 1)}}.
\]

Without explicitly solving \(g^{-1}(I_k)\), we obtain the optimal \(P^*_k\) as
\[
P^*_k = \frac{2(e^{I^*_k} - 1)}{(\phi_{k,1} + \phi_{k,2}) + \sqrt{(\phi_{k,1} + \phi_{k,2})^2 + 4\phi_{k,1}\phi_{k,2}(e^{I^*_k} - 1)}} \quad \forall k,
\]
with \(I^*_k = \log \left(2\phi_{k,1}\phi_{k,2}\lambda^* + \sqrt{(2\phi_{k,1}\phi_{k,2}\lambda^*)^2 + (\phi_{k,1} - \phi_{k,2})^2}\lambda^* \right), \) if \(\lambda^*_k > \frac{1}{\phi_{k,1} + \phi_{k,2}}\);
and \(P^*_k = 0, I^*_k = 0, \) if \(\lambda^*_k \leq \frac{1}{\phi_{k,1} + \phi_{k,2}}\).

Again, \(\lambda^*\) is further determined by the rate condition (16).

### 3.2.3 Scalar Distributed DCM Algorithm

The scalar distributed DCM is referred as the following formulation

\[
\min_{\{I_k\}} NK \cdot P(\{I_k\}), \quad \text{(17)}
\]

subject to \(\sum_{k=1}^{K} I_k \geq R,\)
\(\sum_{k=1}^{K} I_k \geq 0, \quad \forall k,\)
\(\text{(18)}\)
\(\text{(19)}\)

where only a scalar coefficient \(P\) is employed to control the transmission power; i.e., equal power at all spatial sub-channels and all OFDM subcarriers. The optimal value \(P^*\) is unique, as the problem in (17) is convex, and can be found by one-dimensional root-search of \(\sum_{k=1}^{K} I_k(P^*) = R\).

By now, we have presented three distributed DCM algorithms, with different granularity in managing the sub-channels. All three DCM problems are convex optimization problems, which result in tractable and robust numerical solutions, since in convex optimization a local minimum is also a global minimum. Either one-dimensional or scalar distributed DCM only requires feedback of low-rate DCM parameters \(\{P^*_k\}_k\) or \(P^*\) to its own base station; in comparison, the two-dimensional distributed DCM also requires the ideal channel information at the transmitter (of the base station), and it provides the finest sub-channel management.
4 Numerical Results

In this section, we provide numerical results to demonstrate the performance of three distributed DCM schemes proposed in Section 3.2. In the simulations, we adhere to the signal model as in (1); and detailed parameters are set as follows. There are $K = 128$ sub-carriers in the downlink OFDM channels; a proper guard interval is inserted to combat the effect of inter-symbol interference; the time-domain multipath channels consist of six equal-power paths with adjacent delays of highest channel resolution (i.e., the inverse of channel bandwidth). The propagation path loss exponent is $\gamma = 3.8$; the log-normal shadow standard deviation is $\sigma = 8$dB. We adopt a hexagonal layout of total $J = 37$ cells. For the antenna sectoring, we assume three sectors per cell, each using idealized antennas with $120^\circ$ beamwidth and a 20dB front-to-back ratio. (Note that the use of sectoring antennas may compromise the full-scattering assumption of MIMO channels [24], the corresponding data rate difference however is not to be considered in this work.) The performance is shown by the average transmission power per cell versus the requested data rate; for simplicity, we assume that the data rate requests of different users are the same, and having in mind that the proposed DCM schemes can naturally take advantage of the differentiated data rates among users. In each Monte Carlo simulation example, the geographical locations of all MS’s are generated at the beginning and fixed thereafter, where each MS is uniformly placed on a circle with a distance of $d$ from its own BS, and $0 < d < 1$ is normalized by the effective radius of hexagon cell [25]; once again for simplicity, the distance parameter $d$ is the same in all $J$ cells; and the demonstrated performance is the average over 100 random fading channel realizations (i.e., the $h_{i,j,l}$ and $\beta_{i,j}$). Finally, it is assumed that the median of the mean path gain at a reference distance $d = 1$ is one; and it would be better to read the results in multi-cell in a relative sense.

4.1 Performance in Single-cell

In order to gain some insights, before getting into the multi-cell systems, we start with some special examples in single-cell systems, for which the distributed DCM schemes happen to be also global optimal, as $J = 1$. In other words, the single-cell can serve as a benchmark to compare with the multi-cell — since there is no inter-cell interference.
Example 1: Effect of Multiple Antennas

The performance of three DCM schemes is demonstrated for systems with $1 \times 1$ and $4 \times 4$ antenna configurations in Fig 2-3. For both cases, it is seen that the two-dimensional DCM leads to the best performance (i.e., the lowest transmission power); the one-dimensional DCM ranks the second; and the scalar DCM follows as the third. The performance gap between the two-dimensional DCM and the one-dimensional DCM is larger in low data-rate region due to the transmitter precoding (although this precoding gain eventually disappears when data rate goes to infinity [16]); and that performance gap is larger for systems employing more antennas, since the asymmetry among $\phi_{k,s}$ increases as the dimensionality of spatial sub-channels (i.e., $m$) increases [16, 21]. On the other hand, by comparing Fig 2 and Fig 3, it is seen that the one-dimensional DCM harvests less performance gain over the scalar DCM, when more antennas are used in the system; this is because the variation of available information rates on different sub-carriers is averaged down by multiple spatial sub-channels.

Example 2: Effect of Narrow-Band Jamming

In the above example, we considered the impact of spatial sub-channels on different DCM algorithm. In this subsection, we focus on the frequency sub-channels and in particular consider a case of strong narrow-band jamming (NBJ). Fig 4 shows the results in an OFDM-MIMO system with $4 \times 4$ antennas and with 30% sub-carriers being completely blocked by strong NBJ signals. As opposed to Fig 3, where the performance gain of one-dimensional DCM over scalar DCM is marginal due to the averaging of multi-antenna, Fig 4 clearly shows the advantage of one-dimensional DCM over scalar DCM by intelligently avoiding the transmission on those sub-carriers with NBJ.

By far, the above two examples, though simple, all point to one conclusion, that is, the performance of a particular DCM is mostly determined by the magnitude of the imbalance/asymmetry of sub-channels under the direct control of that DCM scheme. Two-dimensional DCM can take advantage all the sub-channel imbalance in both spatial and frequency; one-dimensional DCM is only capable of taking advantage of the imbalance in the frequency domain.

4.2 Performance in Multi-cell

The simulation results in OFDM-MIMO multi-cell systems are illustrated in Fig 5 and Fig 6, respectively for $1 \times 1$ and $4 \times 4$ MIMO cases. In each figure, a group of performance with different
Figure 2: Performance of DCM scheme, the average power $P$ (dB/cell) versus the data rate $R$ (bit/sec/Hz/cell). Note that in single-antenna systems, two-dimensional DCM is exactly the same as one-dimensional DCM.
Figure 3: Performance of DCM scheme, average power $P$ (dB/cell) versus data rate $R$ (bit/sec/Hz/cell).

distance parameters ($d = 0.05, 0.50, 0.95$) is included. In the scope of cellular systems, it is seen that the performance difference of three DCM algorithms in single-antenna systems is marginal. Furthermore, the multi-antenna (MIMO) is shown to be much more helpful in reducing the transmission power than the single-antenna technique for the same data rate request in cellular systems; in other words, given an average transmission power (e.g., $P = 40$dB) and a BS-MS distance (e.g., $d = 0.95$), the $4 \times 4$ systems support much higher data rate (i.e., $R \approx 3.25$bit/Hz/sec) than that (i.e., $R \approx 1.50$bit/Hz/sec) in $1 \times 1$ systems.

It is seen in Figs 5-6 that the average power of the users with distance $d = 0.95$ from the base station practically goes to infinity beyond certain data rate $R$, which can be regarded as the maximum supportable rate for that distance; note that similar thresholds also exist for the cases of $d = 0.05, 0.50$ but correspond to higher data rates (not included in the figures). Moreover, by comparing between Fig 3 and Fig 6 and ignoring the offset due to path-loss, it can be seen that under a wide range of condition (e.g., $R < 2$ bits/Hz/sec) the multi-cell MIMO OFDM systems perform close to the single-cell systems; and the impact of inter-cell interference becomes noticeable at a fairly large data rate (e.g., $R \geq 2.75$ bits/Hz/sec). Regarding other comparisons among the
three distributed DCM schemes, similar observations can be made as in Section 4.1 and are not repeated here.

5 Conclusions

In this work, we have studied dynamic channel management (DCM), more specifically power control and channel allocation, in multi-carrier multi-antenna wireless cellular systems. The receivers feed back the DCM parameters to the corresponding transmitters in order to minimize the transmission power and strive to satisfy every individual user’s data-rate request, on a packet-by-packet basis. In particular, we considered distributed DCM, in which each individual cell locally optimizes its own performance based on the local channel information. By viewing the sub-channels with different granularity, we have derived three distributed DCM schemes, namely the \textit{two-dimensional DCM}, the \textit{one-dimensional DCM} and the \textit{scalar DCM}. The two-dimensional DCM provides the finest channel management; from numerical simulations, its performance gain over the other two DCM schemes is larger especially at lower data rate region and/or with the use of more antennas. Moreover, the
Figure 5: Performance of DCM scheme, average power $P$ (dB/cell) versus data rate $R$ (bit/sec/Hz/cell).

Figure 6: Performance of DCM scheme, average power $P$ (dB/cell) versus data rate $R$ (bit/sec/Hz/cell).
performance gain of the one-dimensional DCM over the scalar DCM is larger when the sub-carrier channels are more asymmetric; but that gain is considerably smaller when more antennas are used. The proposed DCM schemes were simulated in multi-cell systems with realistic parameters; and it is seen that due to the multiple antennas, the maximum supportable rate (i.e., the rate supported by practically finite transmission power) is significantly increased, e.g., by a factor of two for users close to cell boundary. From an overall system perspective, we have also demonstrated that under a wide range of conditions, a multi-cell MIMO OFDM cellular system can achieve almost the same capacity per cell as a single-cell MIMO OFDM system.

6 Appendix

6.1 Optimal solution to one-dimensional distributed DCM in (11)

Lemma 1. The optimization problem in (11) is a convex optimization problem, and its optimal solution is obtained from solving the equations (14)-(16).

Proof. To show the convexity of (11), we need to show that the objective function is a concave function of \( \{P_k\}_k \) and the constraint is a convex set in \( \{P_k\}_k \). First, the objective function (11) is linear in \( \{P_k\}_k \); and a linear function is a concave function. Moreover, the constraint (12)-(13) is a convex set; since the linear combination of convex functions \( I_k(P_k), \forall k \), is a convex function, and the epigraph of a convex function (i.e., the set in the form of (12)-(13)) is a convex set [26].

Next, by making use of the Karush-Kuhn-Tucker (KKT) conditions and the special functional structure of \( P_k(I_k) \), we prove the optimal conditions in (14)-(16). Form the Lagrangian of the optimization problem in (11) as

\[
L \left( \{I_k\}_{k=1}^K, \lambda, \{\gamma_k\}_{k=1}^K \right) = P_k(I_k) - \lambda \left( \sum_{k=1}^{K} I_k - R_j \right) - \sum_{k=1}^{K} \gamma_k I_k, \tag{20}
\]

18
where $\lambda$ and $\gamma_k$ are nonnegative Lagrangian multipliers. The KKT conditions are written as

$$\sum_{k=1}^{K} I_k^* \geq R, \quad I_k^* \geq 0, \quad \forall k, \quad (21)$$

$$\frac{dP_k(I_k^*)}{dI_k^*} - \lambda^* - \gamma_k^* = 0, \quad (22)$$

$$\lambda^* \left( \sum_{k=1}^{K} I_k^* - R \right) = 0, \quad (23)$$

$$\gamma_k^* I_k^* = 0, \quad \forall k, \quad (24)$$

where (21)-(22) are the constraints, directly copied from (12)-(13); (23) is obtained by taking derivative of the Lagrangian with respect to $I_k^*$; and (24)-(25) are the so-called complementary slackness conditions.

We start with the KKT condition in (23). When $I_k^* > 0$, from (25), we have $\gamma_k^* = 0$. Hence, (24) is simplified as

$$\frac{dP_k(I_k^*)}{d(I_k^*)} - \lambda^* = 0. \quad (26)$$

When $I_k^* = 0$, the Lagrangian multiplier $\gamma_k^*$ can take any nonnegative value as $\gamma_k^* \geq 0$; and (24) simplifies to

$$\frac{dP_k(I_k^*)}{d(I_k^*)} - \lambda^* = \gamma_k^* \geq 0. \quad (27)$$

From (4), (26) and (27), we next solve the optimal $\lambda^*$. From (4) and by the strict monotonicity of $P_k(I_k)$ in $I_k$, we have

$$\frac{dP_k(I_k^*)}{d(I_k^*)} \left[ \sum_{s=1}^{m} \frac{\phi_{k,s}}{1 + P_k \phi_{k,s}} \right] - 1 = 0. \quad (28)$$

Substitute in the KKT condition (26) and (27), and define a function $h(P_k)$,

$$h(P_k) \triangleq \begin{cases} 
\lambda^* \left[ \sum_{s=1}^{m} \frac{\phi_{k,s}}{1 + P_k \phi_{k,s}} \right] - 1, & \text{if } I_k^* > 0, \\
(\lambda^* + \gamma_k^*) \left[ \sum_{s=1}^{m} \frac{\phi_{k,s}}{1 + P_k \phi_{k,s}} \right] - 1, & \text{if } I_k^* = 0. 
\end{cases} \quad (29)$$

The function $h(P_k)$ monotonically decreases in $P_k$, since its derivative is non-positive

$$\frac{dh(P_k)}{dP_k} = \begin{cases} 
-\lambda^* \left[ \sum_{s=1}^{m} \frac{\phi_{k,s}^2}{(1 + P_k \phi_{k,s})^2} \right] \leq 0, & \text{if } I_k^* > 0, \\
-(\lambda^* + \gamma_k^*) \left[ \sum_{s=1}^{m} \frac{\phi_{k,s}^2}{(1 + P_k \phi_{k,s})^2} \right] \leq 0, & \text{if } I_k^* = 0; 
\end{cases} \quad (30)$$
moreover

\[
h(P_k)|_{P_k=0} = \begin{cases} 
\lambda^* \left[ \sum_{s=1}^{m} \phi_{k,s} \right] - 1, & \text{if } I_k^* > 0, \\
(\lambda^* + \gamma_k^*) \left[ \sum_{s=1}^{m} \phi_{k,s} \right] - 1, & \text{if } I_k^* = 0,
\end{cases}
\] (31)

\[
h(P_k)|_{P_k=+\infty} = -1,
\] (32)

From the definition in (4), it is clear that \( P_k^* > 0 \iff I_k^* > 0 \) and \( P_k^* = 0 \iff I_k^* = 0 \). Therefore, from (28) and (30)-(32), \( P_k^* = 0 \iff \lambda^* = [\sum_{s=1}^{m} \phi_{k,s}]^{-1} \leq [\sum_{s=1}^{m} \phi_{k,s}]^{-1} \) and \( I_k^* = 0 \). Likewise, \( P_k^* > 0 \iff \lambda^* > [\sum_{s=1}^{m} \phi_{k,s}]^{-1} \) and \( I_k^* > 0 \). From (31)-(32), the existence and uniqueness of \( P_k^* \) are evident. Furthermore, it is not difficult to verify that \( dP_k(I_k^*)/dI_k^* \) is a strictly monotonic function, thus its inverse function exists. So far, we proved (14)-(15).

Finally, since for any \( R_j > 0 \), it results in \( \lambda^* > 0 \); the equality in (16) holds from (24).

\[\square\]

### 6.2 Channel allocation parameters \( \mu_{k,s} \) are redundant in distributed DCM problem (6)

Without loss of generality, we focus on the one-dimensional DCM problem in (11). If channel allocation parameters \( \mu_k \) are introduced, the problem in (11) is re-formulated as

\[
\min \sum_{k=1}^{K} P_k(I_k)
\] (33)

subject to

\[
\sum_{k=1}^{K} \mu_k I_k \geq R,
\] (35)

\[\tilde{I}_k \geq 0, \forall k,
\] (36)

\[
\epsilon \leq \mu_k \leq 1, \forall k,
\] (37)

with \( \tilde{I}_k \triangleq \mu_k I_k \).

This is an integer optimization problem, which is typically not analytically tractable. Following the approach in [12], we look at a related optimization problem, by relaxing \( \mu_k \) to be positive numbers between (0, 1].

\[
\min \sum_{k=1}^{K} \mu_k P_k \left( \frac{\tilde{I}_k}{\mu_k} \right)
\] (34)

subject to

\[
\sum_{k=1}^{K} \tilde{I}_k \geq R,
\] (35)

\[\tilde{I}_k \geq 0, \forall k,
\] (36)

\[
\epsilon \leq \mu_k \leq 1, \forall k,
\] (37)

with \( \tilde{I}_k \triangleq \mu_k I_k \).
In (37), an arbitrarily small positive number \( \epsilon \) is introduced in order to avoid the singularity when \( \mu_k = 0 \), as suggested in [8]. And the relaxed problem (34) is the same as the original problem (33) in the limiting sense when \( \epsilon \to 0 \). Note that although by the same notation of \( P_k \), the \( P_k \left( \frac{I_k}{\mu_k} \right) \) in (34) denotes a two-dimensional function of \( \tilde{I}_k \) and \( \mu_k \), whereas the \( P_k(I_k) \) in (11) denotes a one-dimensional function; we will explicitly highlight this difference anywhere confusion may occur.

The optimization problem in (34) is a well-formulated convex optimization problem.

**Lemma 2.** The optimization problem in (34) is a convex optimization problem.

**Proof.** First, it is true that the constraint set in (35)-(37) is convex set, since there are linear constraints on \((\mu_k, \tilde{I}_k)\). Next, we show that the objective function is convex function in \((\mu_k, \tilde{I}_k)\).

The proof follows the lines in [8]. Observe that the objective function (11) is summation of functions \( f(\mu_k, \tilde{I}_k) \triangleq \mu_k P_k \left( \frac{I_k}{\mu_k} \right), \forall k \). Since summation of convex functions is convex, to prove the convexity of the objective, it is only necessary to show that \( f(\mu_k, \tilde{I}_k) \) is convex in \((\mu_k, \tilde{I}_k)\) in the first quadrant.

A two-dimensional function is convex if and only if its restriction to any line is convex [26]. Let \( g(\mu_k) = f(\mu_k, \tilde{I}_k)|_{\tilde{I}_k=a\mu_k+b} = f(\mu_k, a + \frac{b}{\mu_k}) \), then

\[
g(\mu_k) = \mu_k P_k(a + \frac{b}{\mu_k}) \quad (38)
\]

To show the convexity of \( g(\mu_k) \) for \( \mu_k \geq \epsilon \), we take its second derivative

\[
dg(\mu_k)/d\mu_k = P_k(a + \frac{b}{\mu_k}) - \frac{b}{\mu_k} \left[ dP_k(a + \frac{b}{\mu_k})/d(a + \frac{b}{\mu_k}) \right],
\]

\[
d^2g(\mu_k)/d\mu_k^2 = \frac{b^2}{\mu_k^2} \left[ d^2P_k(a + \frac{b}{\mu_k})/d(a + \frac{b}{\mu_k}) \right]^2 = \frac{b^2}{\mu_k^2} \left[ d^2P_k(I_k)/dI_k^2 \right], \quad (39)
\]

where the last equality in (39) is obtained by change of variable. It is clear from (39) that the convexity of \( f(\mu_k, \tilde{I}_k) \) is equivalent to the convexity of the one-dimensional function \( P_k(I_k) \) in \( I_k \), which can be readily verified.

Therefore, \( g(\mu_k) \) is convex when \( \mu_k \geq \epsilon \), \( \mu_k P_k \left( \frac{I_k}{\mu_k} \right) \) is convex in the first quadrant and (11) is convex in \((\tilde{I}_k, \mu_k)\).

**Lemma 3.** The optimal solution to the convex problem in (34) is obtained by solving the equations (14)-(16), and \( \mu_k^* = 1 \).

Lemma 3 states that, without loss of optimality, one can simply set the channel allocation functions \( \mu_k^* \equiv 1 \); in other words, \( \mu_k^*, \forall k \), are redundant parameters.
Proof. The proof also starts with forming the Lagrangian of (34) as

\[ L = \mu_k P_k \left( \frac{\tilde{I}_k}{\mu_k} \right) - \lambda \left( \sum_{k=1}^{K} I_k - \bar{R} \right) + \sum_{k=1}^{K} \alpha_k (\mu_k - 1) - \sum_{k=1}^{K} \beta_k (\mu_k - \epsilon) - \sum_{k=1}^{K} \gamma_k \tilde{I}_k, \]  

(40)

where \( \lambda, \alpha_k, \beta_k \) and \( \gamma_k \) are nonnegative Lagrangian multipliers. The full set KKT conditions consist of those in (21)-(25), with \( I_k^* \) being changed to \( \tilde{I}_k^* \) and the derivative \( \frac{dP_k(I_k)}{dI_k} \) being replaced by partial derivative \( \frac{\partial P_k}{\partial \tilde{I}_k} \), and the following new conditions involving \( \mu_k^* \) as

\[ 1 \geq \frac{\mu_k^*}{\mu_k^*} \geq \epsilon, \quad \forall k, \]  

(41)

\[ \frac{\tilde{I}_k^*}{\mu_k^*} - \frac{\tilde{I}_k^*}{\mu_k^*} \frac{\partial P_k}{\partial \tilde{I}_k^*} \Bigg| \frac{\partial \tilde{I}_k^*}{\partial \tilde{I}_k} = 0, \]  

(42)

\[ \alpha_k^* (\mu_k^* - 1) = 0, \quad \forall k, \]  

(43)

\[ \beta_k^* (\mu_k^* - \epsilon) = 0, \quad \forall k, \]  

(44)

Besides going through the same steps as in Lemma 1, we need to further show that the optimal \( I_k^* \) is not a function of \( \mu_k^* \), in fact \( \mu_k^* \) takes constant value of 1.

Consider the KKT condition in (42), which is the derivative of Lagrangian with respect to \( \mu_k^* \).

Define a function \( q(\tilde{I}_k^*, \mu_k^*) \), which is affine to the left hand side of (42),

\[ q(\tilde{I}_k^*, \mu_k^*) = P_k \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right) - \frac{\tilde{I}_k^*}{\mu_k^*} \frac{\partial P_k}{\partial \tilde{I}_k^*} \Bigg| \frac{\partial \tilde{I}_k^*}{\partial \tilde{I}_k} \]  

(45)

Taking derivative of \( q(\tilde{I}_k^*, \mu_k^*) \) with respect to \( \frac{\tilde{I}_k^*}{\mu_k^*} \), we have

\[ \partial q(\tilde{I}_k^*, \mu_k^*) / \partial \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right) = \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right) \cdot \frac{\partial^2 P_k}{\partial \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right)^2} \]  

(46)

Moreover, from the definition in (4), it can be shown

\[ \partial^2 P_k \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right) / \partial \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right)^2 = \frac{\sum_{s=1}^{m} \phi_k^s}{(1 + P_k \phi_k^s)^2} > 0. \]  

(47)

From (46) and (47), it is clear that

\[ \partial q(\tilde{I}_k^*, \mu_k^*) / \partial \left( \frac{\tilde{I}_k^*}{\mu_k^*} \right) \begin{cases} < 0, & \text{if } \frac{\tilde{I}_k^*}{\mu_k^*} > 0, \\ = 0, & \text{if } \frac{\tilde{I}_k^*}{\mu_k^*} = 0; \end{cases} \]  

(48)
which indicates that \( q(I_k^*, \mu_k^*) \) is monotonically decreasing in \( \frac{I_k^*}{\mu_k^*} \), and achieves the maximum at \( \frac{I_k^*}{\mu_k^*} = 0 \). That is,

\[
\max \left\{ q(I_k^*, \mu_k^*) \right\} = 0, \quad \text{when } \frac{I_k^*}{\mu_k^*} = 0. \tag{49}
\]

If \( \mu_k^* = \epsilon \), from (43) and (44), it follows that \( \alpha_k^* = 0 \) and \( \beta_k^* \geq 0 \); and (42) is simplified as

\[
q(I_k^*, \mu_k^*) = \beta_k^* = 0. \tag{50}
\]

By (49), (50) holds if and only if \( \frac{I_k^*}{\mu_k^*} = 0 \) and \( \beta_k^* = 0 \). If \( \epsilon < \mu_k^* < 1 \), then \( \alpha_k^* = 0 \) and \( \beta_k^* = 0 \),

\[
q(I_k^*, \mu_k^*) = 0; \tag{51}
\]

which by (49) requires that \( \frac{I_k^*}{\mu_k^*} = 0 \). If \( \epsilon = 1 \), then \( \alpha_k^* \geq 0 \) and \( \beta_k^* = 0 \),

\[
q(I_k^*, \mu_k^*) + \alpha_k^* = 0; \tag{52}
\]

which is valid for \( \frac{I_k^*}{\mu_k^*} \geq 0 \). In summary,

\[
\epsilon \leq \mu_k^* < 1 \iff \frac{I_k^*}{\mu_k^*} = 0, \tag{53}
\]

\[
\mu_k^* = 1 \iff \frac{I_k^*}{\mu_k^*} \geq 0. \tag{54}
\]

The key in (53)-(54) is that the optimal value of \( \mu_k^* \) is not unique when \( \frac{I_k^*}{\mu_k^*} = 0 \). Because when \( \frac{I_k^*}{\mu_k^*} = 0 \), it is sufficient and necessary that \( \lambda^* \leq \left( \sum_{s=1}^{m} \phi_{k,s} \right)^{-1} \), for any value of \( \mu_k^* \), \( \epsilon \leq \mu_k^* \leq 1 \).

Intuitively, the role of \( \mu_k \) is an on-off switch, controlling whether the \( k \)-th channel is active or not; however such an on-off function is redundant with the “water-filling” condition (15), when the \( k \)-th channel is shut off. By letting \( \mu_k^* = 1 \), \( \forall k \), \( \frac{I_k^*}{\mu_k^*} \) is no longer a function of \( \mu_k^* \); and by change of variable, \( \frac{I_k^*}{\mu_k^*} \equiv I_k^* \).

\[\square\]

**References**


